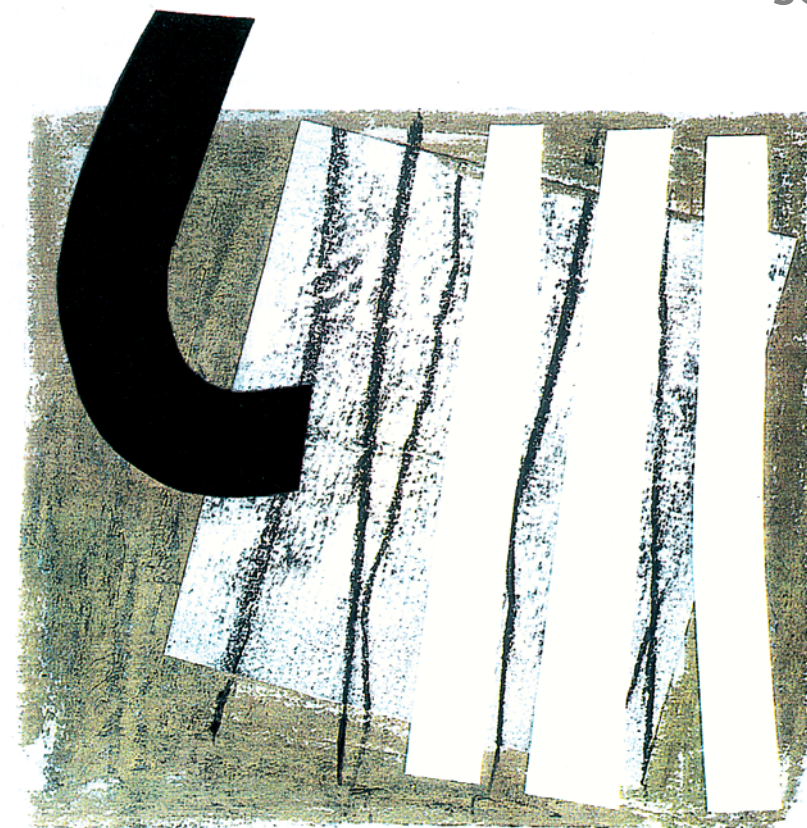




UNIVERSIDAD CARLOS III DE MADRID

## Herramienta didáctica para análisis semántico y traducción de lenguajes formales



documentos  
de trabajo

Prof. Dr. D. Jesús García Herrero  
Prof. Dr. D. Antonio Berlanga de Jesús



UNIVERSIDAD CARLOS III DE MADRID

<http://www.uc3m.es/documentosdetrabajo>  
Serie de Innovación Docente  
nº 06-02-01  
2006





UNIVERSIDAD CARLOS III DE MADRID

## **HERRAMIENTA DIDÁCTICA PARA ANÁLISIS SEMÁNTICO Y TRADUCCIÓN DE LENGUAJES FORMALES**

*Jesús García, Antonio Berlanga*  
*Dpto. Informática, Universidad Carlos III de Madrid*  
*Escuela Politécnica Superior. Avda. Universidad Carlos III 22,*  
*28270 Colmenarejo (Madrid)*

# Herramienta didáctica para análisis semántico y traducción de lenguajes formales

**Jesús García, Antonio Berlanga**

Dpto. Informática, Universidad Carlos III de Madrid  
Escuela Politécnica Superior. Avda. Universidad Carlos III 22,  
28270 Colmenarejo (Madrid)

## 1. Resumen

En este documento se describe una experiencia de innovación docente centrada en el desarrollo de habilidades del alumno (autoaprendizaje y trabajo en equipo) a través de una herramienta informática, que toma parte además en la metodología docente y en el sistema de evaluación del curso.

El proyecto se centró en las asignaturas de la troncalidad de ‘Procesadores del Lenguaje’ de la titulación de Ingeniería Informática de la Universidad Carlos III de Madrid. Su propósito inicial consistió en la mejora del proceso de asimilación de conceptos mediante el autoaprendizaje, con la realización de problemas y prácticas en una herramienta informática con interfaz web disponible para los alumnos matriculados. Esta iniciativa ha fomentado el seguimiento de la asignatura y asimilación de contenidos a través de la realización continua de actividades durante el curso y la realimentación de las correcciones de modo automático. A la vista de los resultados y del cambio de actitud mostrado por los alumnos con respecto a otros cursos, las conclusiones son positivas acerca de su puesta en marcha y continuación en cursos sucesivos.

## 2. Introducción y motivación

La experiencia de innovación docente ha consistido básicamente en el desarrollo y utilización de una herramienta informática para fomentar el desarrollo de habilidades del alumno (autoaprendizaje y trabajo en equipo) y seguimiento continuo de la asignatura, lo que ha tenido impacto además en la metodología docente y en el sistema de evaluación. El proyecto se centró en las asignaturas de ‘Procesadores del Lenguaje 1’ y ‘Procesadores del Lenguaje 2’ de la titulación de Ingeniería Informática de la Universidad Carlos III de Madrid.

Realmente, el documento resume las experiencias correspondientes a dos proyectos de innovación docente, que suponen dos fases de un programa común, con los objetivos mencionados más arriba, particularizados a diferentes aspectos de contenido de las asignaturas. En particular, los proyectos han sido:

- “Desarrollo de herramienta interactiva de análisis sintáctico”. 1a Convocatoria de Apoyo a Experiencias de Innovación Docente, UC3M. Curso 03-04.
- “Herramienta didáctica para análisis semántico y traducción de lenguajes formales”. 2a Convocatoria de Apoyo a Experiencias de Innovación Docente, UC3M. Curso 2004-05

El contexto de esta materia troncal de Procesadores del Lenguaje (12 créditos) de la titulación de Ingeniería Informática es el de dos asignaturas cuatrimestrales de 6 créditos cada una, Procesadores del Lenguaje I y II, que se imparten en el primer y segundo cuatrimestre, respectivamente, del primer año del segundo ciclo. Actualmente hay cuatro grupos (dos en el campus de Leganés, uno en Colmenarejo, y uno en el grupo bilingüe). Los 12 créditos de la troncalidad se reparten a partes iguales entre teoría y prácticas. La materia tiene una fuerte fundamentación teórica, centrada en el conjunto de técnicas formales que soportan los sistemas de traducción de lenguajes formales, análisis de lenguajes de programación y aspectos de implementación de este tipo de sistemas complejos [1-4]. Por otra parte, la componente práctica es muy importante, destacando la familiarización con herramientas para la generación automática de analizadores (Lex/Flex, Yacc/Bison, Jflex/Jcup, JavaCC, etc. ) [1,2,5,6], que son las que permiten abordar un proyecto de estas características de forma realista y, por tanto, es conveniente que sean conocidas por los alumnos.

Posiblemente debido a esta doble naturaleza de fuerte fundamentación teórica y, al mismo tiempo, enfoque práctico para la construcción de sistemas que resuelven unos problemas específicos, se ha detectado que un número importante de alumnos abandona la asignatura sin presentarse a las convocatorias de exámenes, principalmente debido a que la materia teórica y el trabajo práctico deben dominarse simultáneamente. En las clases teóricas se plantean algunos ejemplos y ejercicios ilustrativos, pero es conveniente y necesario que los alumnos realicen ejercicios para la comprensión y asimilación completa de la teoría. Por ello, se planteó la conveniencia de implicar al alumno en la materia fomentando el autoaprendizaje. Con este enfoque, la herramienta interactiva propuesta se plantea para ayudar al alumno en la realización de la práctica obligatoria y para agilizar y mejorar el aprendizaje de la materia teórica. En las clases teóricas se plantean algunos ejemplos y ejercicios introductorios, pero es muy conveniente que los alumnos realicen ejercicios complementarios para la comprensión completa de la materia.

### 3. Objetivos

El objetivo fundamental ha sido el de aportar una herramienta que facilite a los alumnos la comprensión de los conceptos teóricos de la materia de traducción y compilación de lenguajes, a través de su aplicación a ejemplos ilustrativos. Más concretamente, la puesta en marcha de esta herramienta persigue los siguientes objetivos, particularizados para el alumno y el profesor.

Para el alumno:

- Fomentar el trabajo personal, involucrando al alumno en la tarea de aprendizaje y seguimiento continuado de la asignatura.
- Tomar conciencia del conocimiento real que se tiene acerca de la materia de la asignatura.
- Explorar diferentes alternativas de solución, pudiendo razonar y posteriormente verificar la validez de estas.
- Fomentar la discusión y debate entre los alumnos acerca de soluciones y estrategias a ciertos tipos de problemas planteados en el contexto de la materia impartida.

Para el profesor:

- Disponer de medios para proponer y evaluar actividades a lo largo de todo el cuatrimestre que implique al alumno en la asignatura desde el principio.



- Agilizar cálculos mecánicos y poder abordar en clase un número mayor de ejemplos.
- Mostrar otros aspectos importantes, como la visualización paso a paso del proceso de análisis.
- Realimentación de la resolución de los ejercicios: detección del grado de seguimiento, errores frecuentes, carencias específicas, etc.

Estos objetivos impuestos a la iniciativa aquí presentada han perseguido una orientación docente hacia competencias y habilidades deseables en los alumnos que cursen la asignatura, especialmente la capacidad de razonar y relacionar los contenidos teóricos de la materia y la capacidad de soltura en su aplicación a problemas prácticos. Así mismo, los objetivos van en consonancia con las tendencias apuntadas en la declaración de Bolonia [7,8]: distribución del esfuerzo dedicado a la superación de la asignatura a lo largo de todo el cuatrimestre, incrementar la dedicación de los alumnos a actividades fuera de las aulas, reduciendo parte de las clases magistrales para incrementar los debates y resolución de casos prácticos, así como incrementar el peso dado a las actividades prácticas.

#### **4. Metodología**

La metodología que se ha llevado a cabo para integrar esta herramienta en la docencia ha sido la propuesta y entrega de diferentes problemas propuestos por el profesor que los alumnos deberán analizar y resolver. Los alumnos utilizan la herramienta para el análisis y evaluación de las soluciones propuestas antes de su entrega. Las actividades propuestas pueden ser individualizados para distintos grupos de alumnos con el fin de mantener variedad de ejercicios y evitar soluciones compartidas por diferentes grupos.

En la herramienta, además de servir de autoevaluación, se registran los pasos dados por los alumnos para proponer sus soluciones y modificarlas. Esto permitió valorar los puntos oscuros de la materia y el progreso en el aprendizaje, a través de los errores cometidos sistemáticamente en determinados temas y del número de intentos realizados antes de finalizar correctamente un ejercicio. De esta manera, la interacción con la herramienta de cada alumno queda almacenada para su posterior análisis (alumnos que proponen soluciones a los ejercicios, número de intentos, nivel de fracaso, etc.).

Por otro lado, se mantiene abierto un "foro electrónico", de manera que cada grupo tendrá que aportar unas conclusiones y resumen de su trabajo, sobre el que otros grupos podrán realizar preguntas y comentarios sobre determinados aspectos acerca de los ejercicios realizados.

Un aspecto destacable con respecto a la metodología ha sido el hecho de mantener coordinados todos los grupos en los que se ha aplicado (4 grupos). Para ello se encargaron dos profesores de la elaboración de actividades, sesiones comunes introductorias en los diferentes grupos, y de la configuración y mantenimiento del servidor para tener todas las actividades compartidas entre los diferentes grupos, manteniendo plazos de entrega comunes.

#### **5. Descripción de la herramienta desarrollada en el proyecto**

La herramienta desarrollada en este proyecto consiste básicamente en un meta-compilador interactivo. Facilita la visualización detallada los pasos de construcción y de utilización de un compilador/traductor de lenguajes formales, añadiendo la lógica necesaria para que

sirva como herramienta de autoevaluación de los ejercicios propuestos a lo largo del curso. El formato es el de un programa de tipo “applet” de Java, de manera que se abre con un navegador web, con acceso restringido a los alumnos registrados y profesores de las asignaturas. Por tanto, la ejecución del programa es local a la máquina utilizada por el alumno, si bien los resultados son enviados de vuelta al servidor. De este modo, al estar centralizada la recogida de datos de utilización, se puede aplicar con facilidad herramientas estadísticas para el análisis de participación y evolución durante el curso. Dada la extensión de las diferentes técnicas existentes, el desarrollo de esta herramienta se realizó en dos fases, como se indicó en la introducción. En la primera fase, se abordó el desarrollo de la herramienta interactiva de análisis sintáctico, que es el motor de todo proceso de traducción, y la puesta en marcha del servidor. En una segunda fase se realizó el desarrollo de las etapas posteriores de análisis semántico y generación de código así como el funcionamiento paso a paso.

### 5.1 Descripción

Para el análisis sintáctico se toma la definición de una gramática formal en formato BNF y genera las tablas de análisis sintáctico: LL(1), SLR(1), LR(1) y LALR(1), calculando, además, los conjunto y estados internos empleados para construir las tablas [1]. En el caso de que la gramática no sea de la clase contemplada por cada uno de los analizadores sintácticos implementados, que son los que cubre la materia objeto de interés, entonces muestra las entradas en las tablas que se producen las ambigüedades.

En la siguiente figura (puede visualizarse en la dirección <http://www.giaa.inf.uc3m.es>) se muestran, tras la validación del usuario, las diferentes funciones que pueden realizarse, como pestañas de selección. La interfaz proporciona facilidades para definir de diferentes modos la gramática que se utilizará para el análisis (propuesta en el servidor por parte del profesor, carga de fichero local y edición en la propia herramienta).



Figura 1. Funcionalidades de la herramienta de análisis

Una vez que la gramática está definida, a partir de este punto puede abordarse la obtención de las diferentes tablas de análisis implementadas, que son las que cubre la materia de Procesadores del Lenguajes para la que se ha desarrollado: conjuntos PRIMERO y SIGUIENTE, en la pestaña “**Conjuntos**”; las diferentes tablas posibles, en la pestaña

“**Tablas**”; la pestaña “**Estados**” visualiza los estados utilizados en la construcción de las diferentes tablas, etc. En la siguiente figura se muestra la representación que se realiza en el sistema para cada uno de los pasos anteriores.



Figura 2. Elementos del análisis sintáctico

La última funcionalidad, “**Análisis**”, permite visualizar los diferentes pasos de análisis de un texto de entrada, previamente preprocesado en tokens, con la presentación de los resultados obtenidos en un análisis semántico (valores de atributos y ejecución de acciones). En la figura siguiente se muestra los elementos incorporados para esta funcionalidad. Dada la envergadura de esta parte del análisis, su implementación se llevó a cabo una segunda edición de los proyectos de innovación docente mencionados.



Figura 3. Análisis semántico y visualización “paso a paso”

Para incluir acciones semánticas en la gramática, pueden declararse acciones al final de cada producción (esquema de evaluación de atributos “sintetizados” [1]). El formato para declarar una acción semántica es:

[ \$0 es "valor1" concat "\$1" concat "valor2" ]

Donde cada palabra debe ir separada por espacios. A efectos ilustrativos y con objeto de poder realizar numerosos ejemplos didácticos, los atributos y acciones contemplados son los correspondientes a una manipulación de valores textuales, que permiten mostrar ejemplos de traducción de lenguajes:

- asignar.- Mediante la palabra reservada "es".
- concatenar.- Mediante la palabra reservada "concat".

Además, esta función permite ejecutar acciones semánticas sobre una sentencia de entrada de análisis, que es introducida en el programa por fichero siguiendo un formato predeterminado. Una vez introducida la cadena de tokens a analizar y elegido el algoritmo a utilizar, podemos comenzar con el análisis semántico. Las dos posibilidades disponibles son realizar un análisis completo de una sola vez, en cuyo caso se muestra si la cadena de entrada es aceptada o no y el resultado, o realizar un análisis paso a paso, lo que permite observar el funcionamiento de cada uno de los algoritmos. Para su ejecución nos apoyamos en la utilización de una pila, que aparece en la pantalla de la izquierda (figura 3), y a través de los botones “Siguiente” y “Anterior”, que nos permitirán ir hacia delante o hacia atrás en la ejecución del algoritmo y visualizar el efecto de cada una de las acciones semánticas indicadas en la definición de la gramática.

Al final de la ejecución, al fondo de la pila semántica podemos observar el resultado de la ejecución de las diferentes acciones, así como en la ventana “log” del programa.

## 5.2. Modos de utilización

La herramienta interactiva funciona en modo de “*applet*” en una página web a la que se accede previo registro por los alumnos. Los profesores de la asignatura proponen problemas de teoría de análisis sintáctico: obtención de una gramática G2 que reconozca un lenguaje especificado, resolución de ambigüedades, etc. La configuración del modo de uso de la herramienta (edición o presentación de resultados) se realiza en lenguaje HTML



y determina el tipo de ejercicio propuesto. A través del registro, la herramienta devuelve a los profesores información acerca de los pasos dados por los alumnos para resolver los problemas propuestos. Se han implementado las opciones de utilización que se describen a continuación, con objeto de poder proponer diferentes tipos de actividades:

### ***A. Ejercicios cerrados: cálculo de conjuntos y tablas para una gramática propuesta***

Este primer modo es el que sirve para ejercitar los primeros conceptos presentados de la asignatura. Aquí se propone una gramática específica (se deshabilita la opción de cargar o editar) y se dejan abiertas las soluciones para los conjuntos y tablas, de manera que el alumno únicamente puede proponer los valores correspondientes y comprobar la corrección. El alumno debe editar los conjuntos y tablas y podrá ir comprobando la solución que propone hasta dar con la correcta. Todo el proceso se traza en un archivo en el servidor donde se podrán comprobar errores e intentos (el nombre del archivo se puede especificar por parámetro). Los resultados propuestos se envían al servidor en el momento de presionar el botón de comprobación.



Figura 4. Ejemplo de edición de conjuntos

### ***B. Problemas abiertos: edición y análisis de gramáticas***

Este segundo modo sirve para que el alumno, con un conocimiento mayor acerca de los mecanismos de análisis sintáctico, pueda experimentar con sus propias gramáticas y analizar sus propiedades para enfrentarse a un problema abierto (diseño de lenguaje, propiedades de gramáticas, exploración de casos posibles, etc.). El alumno puede editar una gramática (cargándola de fichero o sobre la propia pantalla) e ir comprobando resultados, y cuando este conforme enviar esa gramática al servidor. También la puede guardar localmente para seguir en otra sesión.

## 6. Resultados

### 6.1. Resultados en el servidor

Como se indicó en la introducción, la herramienta desarrollada se instala en un servidor web de manera que recoge toda la información de los alumnos registrados que acceden, a través de su número de identificación:



Figura 5. Almacenamiento de datos de los alumnos que acceden

De esta manera, toda la “historia” de interacción de cada alumno con la herramienta queda registrada en su directorio del servidor, con ficheros asociados a la y fecha de acceso y la solución o soluciones propuestas, según el tipo de ejercicio, lo que permitirá realizar un seguimiento individualizado de la utilización y evolución de cada alumno usuario a lo largo del curso. La ventaja para el profesor de esta realimentación que provee la herramienta acerca de la resolución de los ejercicios propuestos es que permite detectar no solo carencias colectivas, sino también aquellas en las que pueden incurrir alumnos específicos.

A modo de ejemplo de análisis, se muestran a continuación el realizado sobre los ejercicios entregados por los alumnos del curso 2004-05. El número total de alumnos en los grupos analizados fue de 130

Tabla 1. Número de alumnos matriculados

GRUPO	Alumnos
COLMENAREJO	7
LEGANES-81	54
LEGANES-82	69
Total	130

Se distingue el número de ejercicios entregados y los correctos. Los promedios de ambas variables se muestran en la tabla siguiente, y la figura indica la distribución detallada de ambas variables.

Tabla 2. Ejercicios entregados y correctos

	GRUPO			Total general
	COLMENAREJO	LEGANES-81	LEGANES-82	
Promedio de usa	100%	67%	72%	72%

Ejercicios entregados (promedio)	12,85	13,58	11,8	12,56
Ejercicios correctos (promedio)	11,14	12,66	10,82	11,55

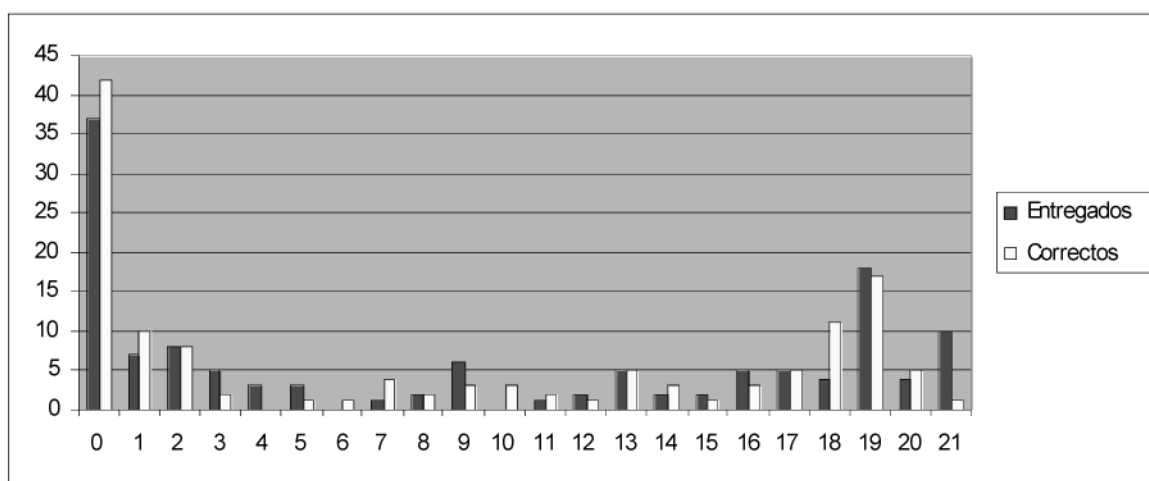


Figura 6. Ejercicios entregados por los alumnos matriculados

A la vista de estos resultados, la participación de los alumnos del grupo fue muy elevada (72%), y observando el histograma pueden distinguir claramente dos tipos, los que abandonaron tras realizar y entregar menos de la mitad de los ejercicios y los que entregaron un número significativo de ejercicios, grupo bastante más numeroso. Es reseñable el elevado número de alumnos que entregaron prácticamente la totalidad de los ejercicios propuestos.

### 6.1. Otros datos

En la siguiente figura se muestra la dedicación total en horas de los alumnos a las asignaturas, distribuidas a lo largo de los temas. Incluye dedicación a contenidos teóricos (incluyendo resolución de actividades) y la realización de la prácticas. Estos resultados muestran la mayor dedicación a aquellos temas cubiertos por actividades complementarias (análisis sintáctico y semántico), destacándose también la dedicación a otras actividades (preparación y presentación de trabajos al final del curso). Lamentablemente no se dispone de datos de dedicación en cursos anteriores, lo que hubiera servido para mostrar el impacto de esta actividad en la dedicación, que presuponemos ha experimentado un notable crecimiento.

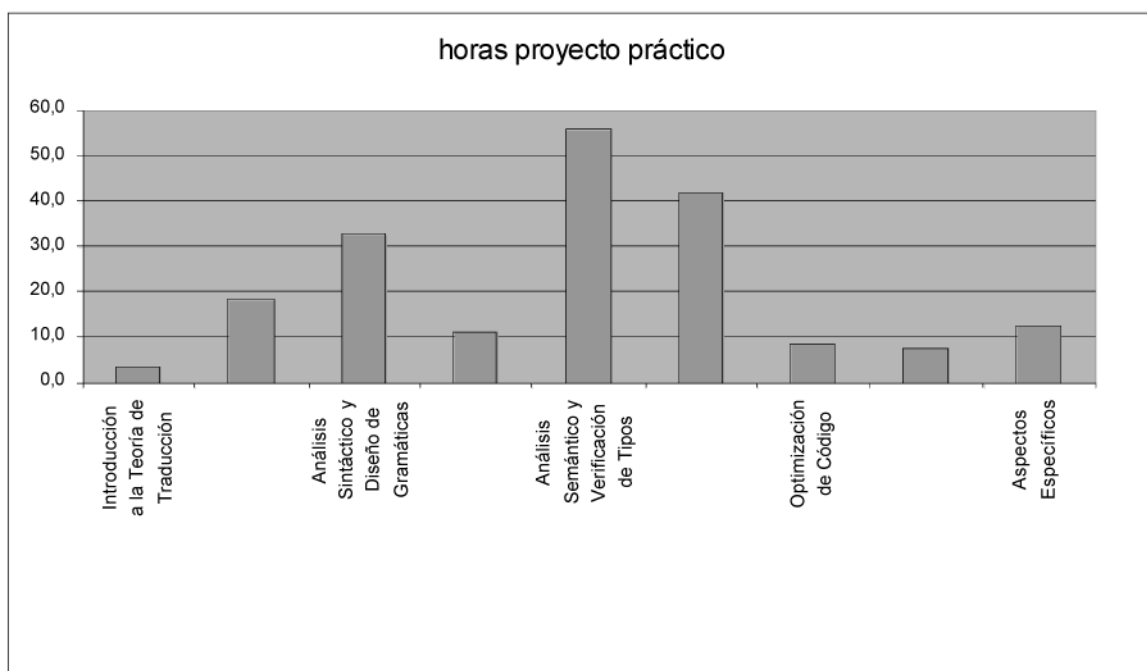
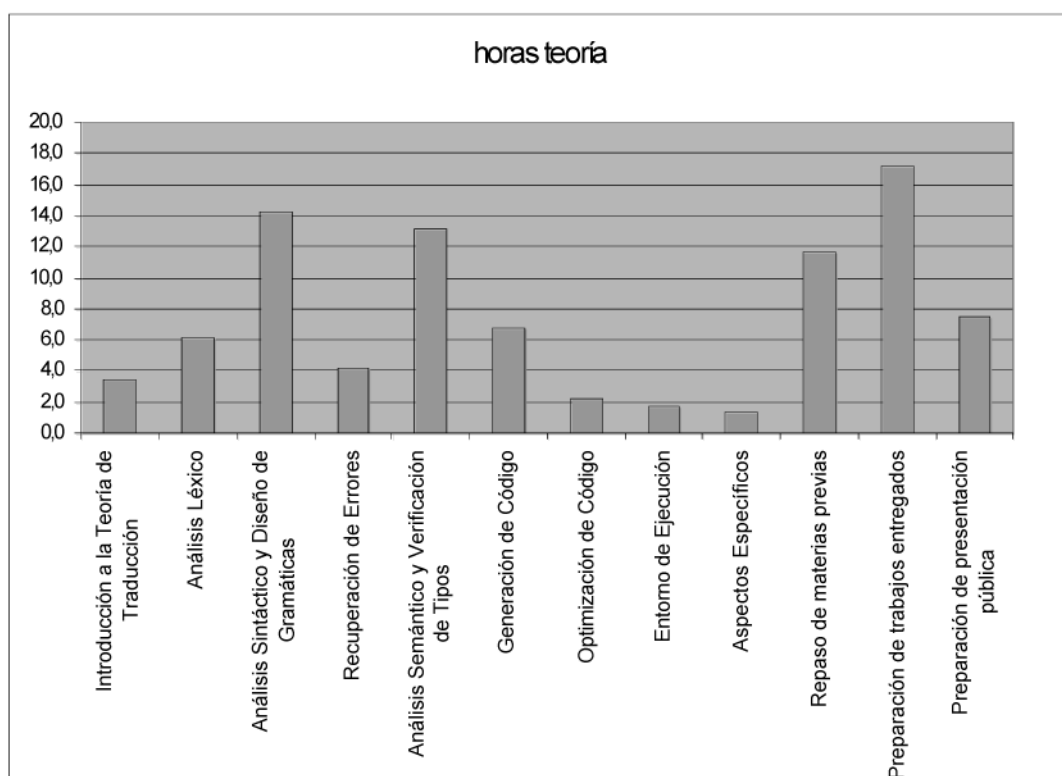


Figura 7. Dedicación de los alumnos a los temas de la asignatura (teoría y práctica)

Tabla 3. Resumen de dedicación total a las asignaturas de Procesadores del Lenguaje

Horas dedicación	Teoría	Práctica
Promedio	179,25	122,1579
Desviación	87,58086	52,03659



Sí disponemos, en cambio, de los resultados académicos resumen de los últimos cursos, mostrados aparecen en la tabla 4. Aquí puede resaltarse la mejora significativa en el número de alumnos que superaron la asignatura con la puesta en marcha de esta experiencia en comparación con otros cursos. En opinión de los profesores, esta mejora es debida a la implicación de un elevado porcentaje de alumnos en el seguimiento continuo de la asignatura y realización en plazo de las actividades propuestas, lo que marca un ritmo de trabajo que permite a los alumnos superar las asignaturas con más facilidad.

Tabla 4. Resultados académicos en los últimos cursos de Procesadores del Lenguaje

Curso	04/05	03/04	02/03
No presentado	19.0%	8.6 %	8.1 %
Libre dispensa	0%	5.7 %	10.8 %
Suspenso	8.6%	37.1 %	35.1 %
Aprobado	39.6	22.9 %	21.6 %
Notable	29.3%	22.9 %	21.6 %
Sobresaliente		3 %	2.7 %
Matrícula de Honor	3.4%		

## 7. Conclusiones

Podemos concluir que el nivel de satisfacción de los objetivos planteados ha sido elevado, a la vista de la opinión de los alumnos y de los resultados. Podemos destacar la alta participación en el sistema nuevo de evaluación con la entrega en plazo de las actividades propuestas a lo largo de todo el cuatrimestre, lo que ha redundado en una notable mejora de los resultados académicos con respecto a otros cursos.

La herramienta desarrollada queda disponible para cursos sucesivos, en los que la aplicación podrá ser más exhaustiva y de mayor aprovechamiento, al corregirse deficiencias observadas y no necesitar tanto tiempo para su desarrollo y puesta en marcha, lo que supuso gran parte de los recursos dedicados en los dos cursos en los que se llevó a cabo la experiencia. Un elemento importante ha sido mantener una elevada coordinación entre todos los grupos en los que se ha llevado a cabo la experiencia, con profesores implicados en la elaboración de actividades, configuración y mantenimiento del sistema.

## Bibliografía

- [1] A. V. Aho and Ravi Sethi and J. D. Ullman, Compiladores: Principios, Técnicas y Herramientas, Addison-Wesley Iberoamericana, 1990.
- [2] Kenneth C. Loudon: "Construcción de Compiladores. Principios y práctica". Thomson, 2004
- [3] F. J. Sanchis and C. Galán, Compiladores: Teoría y Construcción, Paraninfo, 1986.
- [4] Dick Grune, Henri E. Bal, Criel J.H. Jacobs, Koen G. Langendoen: "Modern Compiler Design". John Wiley & Sons, 2000
- [5] Garrido, Iñesta, Moreno, Pérez: "Diseño de Compiladores". Publicaciones Universidad de Alicante, 2002.
- [6] S. Gálvez Rojas, M. A. Mora Mata. Traductores y compiladores con Lex/Yacc, Jflex/Cup y JavaCC. Edición Electrónica. Universidad de Málaga, 2005 MA-185-2005

[7] Joint Declaration of the European Ministers of Education, Bologna 1999.

[8] M. de Miguel Díaz. Metodologías de enseñanza y aprendizaje para el desarrollo de competencias: orientación para el profesorado universitario ante el espacio europeo de educación. Alianza, 2006.

### Agradecimientos

mejora  
docencia  
aprendizaje



Expresamos nuestro agradecimiento al servicio de gestión académica de la universidad por el apoyo prestado y la ayuda concedida a esta iniciativa. A los becarios que participaron en el desarrollo de la experiencia, Simón de Pablo Bruhlmann (Ing. Telecomunicación) y Abel Cal Rodríguez (Ing. Informática). También a los alumnos de la asignatura por su colaboración y sugerencias a lo largo de esta experiencia.







